SWR
Spoken Word Recognition

教材



単語音声の認識　DAY3

参照ページ

第 6 章 単語認識	6-1〜6-36
付録 A 数字単語音声のスペクトル分析実例	A-1〜A-7

自分で発声した単語の音声データを録音し，学習データを作り，
スペクトル分析を行って，単語 HMM の学習を行います．その単語 HMM を用いて単語認識の実験を行います．
可能であれば，個性や創意を発揮し，認識対象単語（5 種類以上）は実習テキストと異なるものを考えてください．

本日の実習手順

§6.5.1，§6.5.2
認識用設定ファイルの作成
単語音声学習データの作成
(録音→ラベル作成→分割→検聴)

↓

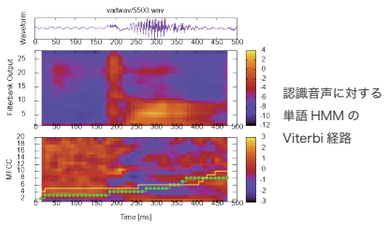
§6.5.3
スペクトル分析

↓

§6.5.4
単語 HMM の学習

↓

§6.5.5
On-The-Fly 単語認識



認識音声に対する
単語 HMM の
Viterbi 経路

学習データの作成には手間がかかる

多数の単語発話にラベルを付けるのは手間がかかります（5 単語の場合，標準作業時間は 90 分）．単語 1 種類あたり 10 個の学習データで十分ですが，少なくとも 12 個の発声を録音してください．その中から学習データとして適当なものを選んでください．　pp.6-10〜23

On-The-Fly 単語認識の結果には音声区間検出の成否が大きく影響

マイク入力の音響信号から音声区間を検出するために，本実習では最も基本的な音声のパワー [dB] の閾値を用いる方法を使っています．声が小さすぎると単語認識が実行されず，大きすぎると単語が途中の無音区間で切り取られてしまうことがあります．落ち着いて，単語 HMM の学習用音声データを発話したときと同じ程度の音量で，同じ発音で発声しましょう．ヘッドセットのマイクロフォンの位置も影響するので，出来るだけいつも同じ位置にセットしてください．　§6.3.2，§6.5.6.3

DAY3　レポート　(課題の詳細は PDF 版テキストを参照のこと)

(1) 単語認識実験
　(a) 単語認識タスクを設計してください .
　　・あなたの単語認識タスクの目的を述べてください．仮に使うとしたらどのような用途に用いますか？
　　・5 種類以上の単語からなる単語集合を定義し，一覧表 (表 6.1 の 形式) を作ること .
　　・各単語のモデルに用いた HMM の状態数を上記一覧表に記述すること .
　(b) HMM の学習曲線 (図 6.5 に相当するもの). 認識対象の単語のうち 任意の 1 つに対するものでよい .
　(c) 学習の検証の結果
　(d) On-The-Fly 単語認識の結果と考察
(2) このテーマについて（第 1 週から第 3 週の総括）
　(a) ポイントは何であったか？　(b) 良くわかったこと
　(c) わからなかったこと　(d) 要望
　(e) 感想，その他

SWR
Spoken Word Recognition

教材

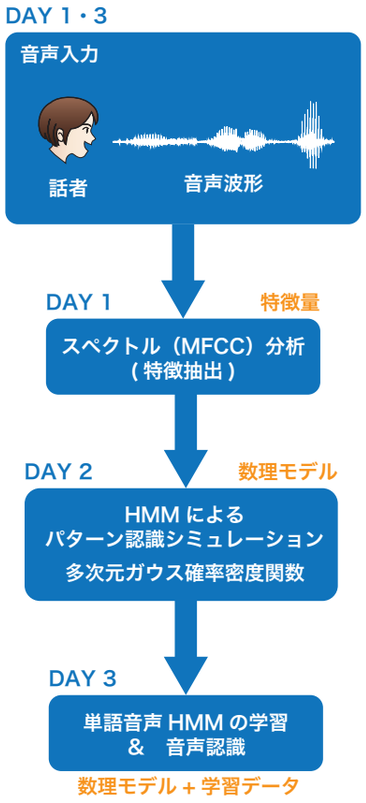


このテーマについて

参照ページ

第 1 章 はじめに	1-1〜1-9
第 7 章 おわりに	7-1〜7-3
関連文献	7-5〜7-7

音声認識は人間の言語発話の内容と意味を機械が認識するための技術です．
メディア情報検索・認識（音声）では，音声認識の最も基本的な形態である**特定話者小語彙孤立単語音声認識**の実験を行います．認識処理に必要な全てのソースコード（C 言語）を提供します．



数理モデル ＋ データ ＝ 機械学習・パターン認識

機械学習とは，明示的な規則を用いることなく，その代わりに学習データと数理（推論）によってある特定の課題（これをタスク（task）と言います）を実行するためのアルゴリズムと統計的モデルに関する技術のことです．機械学習は事例「**学習データ**」に基づいて**数理モデル**のパラメータの値を決め，この数理モデルを用いて認識や予測を行うことができるようにします．

音声認識の場合，音声波形データをスペクトル分析し，音韻をよく表すような**特徴量**に変換したものを学習データとして用います．良い数理モデルを学習して高い認識率を得るためには，学習データは質の揃ったものを多数用意しなければなりません．

実習テキストでは数字音声認識をタスクとして説明しています．数字は種類が少ない割に応用範囲が広いのがその理由です．受講者の皆さんは個性と創意を発揮したタスクを考えてください．

実習テキスト

音声認識に関する基礎知識，数理，実験手順の詳細を解説したテキストは PDF で提供しています．このパンフレット，プレゼン資料とともに活用してください．
ファイル名：**~/asr/studentbook.pdf**
Ubuntu のドキュメント・ビューア：　`% evince ~/asr/studentbook.pdf &`

成績評価・レポート

「メディア情報学実験 (I 類 メディア情報学プログラム) 受講要領」（2026 年 10 月 1 日）を適用します．

レポートは毎回提出（計 3 回）．実験日の翌週の水曜日の 23:59:59 までに asr@takagi.inf.uec.ac.jp に提出すること．
レポートのメール送信は以下のように定めます．
　・ファイル形式：Portable Document Format（PDF）
　　※PDF ファイルに対しセキュリティ設定（ファイルロック，パスワード等）を行わないこと．
　・ファイル名：半角英数字で “学籍番号 _ 通し番号 _ 姓” とすること（例：1234567_1_Takagi.pdf）．
　・メールの Subject は PDF のファイル名と同一すること．

その他は，実験テーマ教材ウェブサイトの**レポート**を参照のこと．

SWR
Spoken Word Recognition

教材



特徴抽出　DAY1

言語音声発声の仕組みと音声認識のためのスペクトル分析法を学びます．
メル周波数ケプストラム係数（MFCC）分析のための，フィルタバンク分析と離散コサイン変換の数値計算をプログラミングします．音声収録・編集・再生のためのツールの使い方を練習し，自分の音声の MFCC 分析を行います．

参照ページ	
第 2 章　音声	2-1〜2-12
第 3 章　スペクトル分析	3-1〜3-24
付録 A　数字単語音声のスペクトル分析実例	A-1〜A-7

本日の実習手順

§1.7
計算機実習の準備
ファイルのコピー ＆ 環境設定の追加
↓
§3.10.1
[実習] フィルタバンク分析
↓
§3.10.2
[実習] MFCC 分析（離散コサイン変換）
↓
§3.10.3-1
[実習] 音声収録の練習
↓
§3.10.3-2
[実習] スペクトル分析

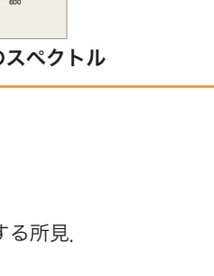
C 言語の整数除算

整数の割算では小数部分は切り捨てられます．これを避けるのであれば，型変換を使用します．

（例）
```
int i, j; i=1; j=2;
float r, s;
r=i/j; //r の値は 0
s=(float)i/(float)j;
        // s の値は 0.5
```

式 (3.2) の実装等で注意．

ヘッドセット



マイクブームの角度を変えるときは根元を回転させること．微調整はブームを直接曲げて行う．急な角度で曲げると断線の原因になる．ヘッドセットは PC 本体前面のコネクタにしっかりと奥まで挿し込むこと．抜くときはプラグ部を持って抜くこと．

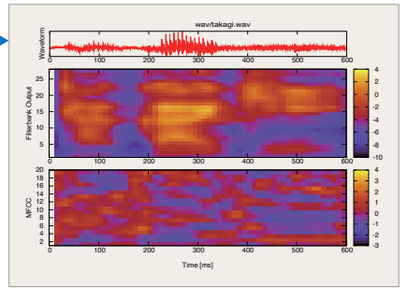


図 3.23　自分の名前（姓または名）のスペクトル

DAY1　レポート　（課題の詳細は PDF 版テキストを参照のこと）

(1) 実習課題の目的
(2) `ad2fb.c` および `fb2mfcc.c` の穴埋め部分と動作検証結果．
(3) 自分の名前の分析結果：`drawspec` のフィルタバンク出力およびそのパターンに関する所見．
(4) 考察
(5) 一般事項　(a) 本日のポイントは何であったか？　(b) 良くわかったこと
(c) わからなかったこと　(d) 要望
(e) 感想，その他

2026

SWR
Spoken Word Recognition

教材



HMM の算法　DAY2

音声パターンの統計的数理モデルとして有名な隠れマルコフモデル（HMM）の定義と算法を学び，最も単純な事例を用い HMM によるパターンの学習と認識の実習を行います．単語音声スペクトルのモデル化に用いる多次元ガウス確率密度関数のプログラミングも行います．

参照ページ	
第 4 章 統計的音声認識	4-1〜4-3
第 5 章 隠れマルコフモデル	5-1〜5-27

本日の実習手順

§5.6.1
[実習] forward アルゴリズム
↓
§5.6.2
[実習] backward アルゴリズム
↓
§5.6.3
[実習] Viterbi アルゴリズム
↓
§5.6.4
[実習] Baum-Welch アルゴリズム
↓
§5.6.4.1・4.3
[実習] HMM の学習実験
↓
§5.6.4.2・4.3
[実習] HMM による認識実験
↓
§5.6.5
[実習] 多次元ガウス確率密度関数

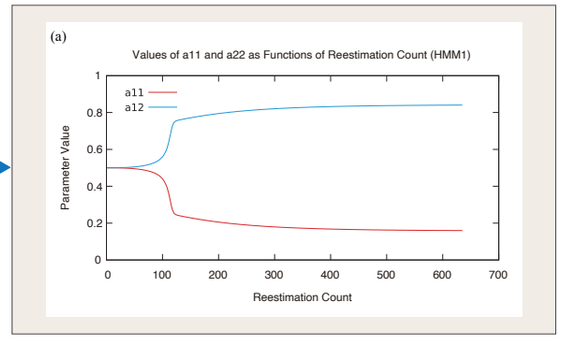


図 5.14(a): HMM_1 の状態遷移確率 a_{11} と a_{12} が再推定の繰り返しにより信号源のパラメータに近づく様子．

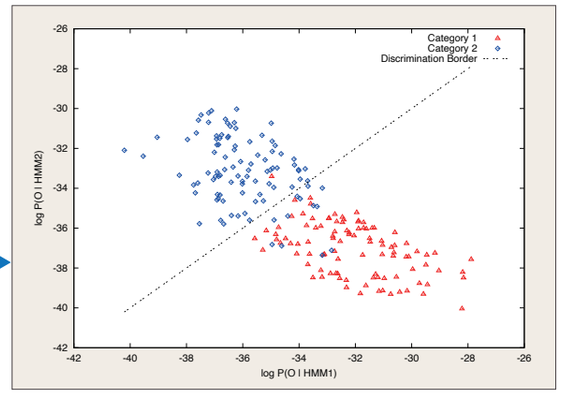


図 5.12　2カテゴリのパターン認識．点線は識別平面．

DAY2　レポート　（課題の詳細は PDF 版テキストを参照のこと）

(1) 実習課題の目的
(2) HMM 算法のプログラミングとシミュレーション
(a) Forward アルゴリズム：`forward.c` の穴埋め部分，`drillF` の実行結果．
(b) Backward アルゴリズム：`backward.c` の穴埋め部分，`drillB` の実行結果．
(c) Viterbi アルゴリズム：`viterbi.c` の穴埋め部分，`drillV` の実行結果．
(d) Baum-Welch アルゴリズム：`baumwelch.c` の穴埋め部分，`drillT` による学習結果と学習曲線，`drillR` による認識結果および `drawR` による識別図．
(e) 多次元ガウス確率密度関数：`gpdf.c` の穴埋め部分，`drillG` の実行結果．
(3) 考察
(4) 一般事項　(a) 本日のポイントは何であったか？　(b) 良くわかったこと

2026